On Goodness-of-fit Testing for Ergodic Diffusion Process with Shift Parameter*

Ilia Negri, Li Zhou†

Abstract

A problem of goodness-of-fit test for ergodic diffusion processes is presented. In the null hypothesis the drift of the diffusion is supposed to be in a parametric form with unknown shift parameter. Two Cramer-Von Mises type test statistics are studied. The first one is based on local time estimator of the invariant density, the second one is based on the empirical distribution function. The unknown parameter is estimated via the maximum likelihood estimator. It is shown that both the limit distributions of the two test statistics do not depend on the unknown parameter, so the distributions of the tests are asymptotically parameter free. Some considerations on the consistency of the proposed tests and some simulation studies are also given.

Keywords: Ergodic diffusion process, goodness-of-fit test, Cramer-Von Mises type test.

1 Introduction

We consider the problem of goodness of fit test for the model of ergodic diffusion process when this process under the null hypothesis belongs to a

*This work has been partially supported by MIUR grant 2009.
†Corresponding author. *E-mail address*: li.zhou.etu@univ-lemans.fr.

Ilia Negri
Department of Information Technology and Mathematical Methods, University of Bergamo. Viale Marconi 5, 24044 Dalmine (BG) Italy.

Li Zhou
Laboratoire de Statistique et Processus, Université du Maine
72085 Le Mans Cedex, France.

given parametric family. We study the Cramer-von Mises type statistics in two different cases. The first one is based on local time estimator and the second one is based on empirical distribution function estimator. We show that the Cramer-von Mises type statistics converge in both cases to some limits which do not depend on the unknown parameter, so the test is asymptotically parameter free (APF).

Let us remind the similar statement of the problem in the well known case of the observations of independent identically distributed random variables $X^n = (X_1, \ldots, X_n)$. Suppose that the distribution of X_j under hypothesis is $F(\vartheta, x) = F(x - \vartheta)$, where ϑ is some unknown parameter. Then the Cramer-von Mises type test is

$$\hat{\psi}_n(X^n) = 1\!\!\mathrm{I}_{\{\omega_n^2 > e_\varepsilon\}}, \qquad \omega_n^2 = n\int_{-\infty}^{\infty}\left[\hat{F}_n(x) - F\left(x - \hat{\vartheta}_n\right)\right]^2 \mathrm{d}F\left(x - \hat{\vartheta}_n\right)$$

where the statistic ω_n^2 under hypothesis converges in distribution to a random variable ω^2 which does not depend on ϑ. Therefore the threshold e_ε can calculated as solution of the equation

$$\mathbf{P}\left\{\omega^2 > e_\varepsilon\right\} = \varepsilon.$$

The details concerning this result can be found in Darling [3]. For more general problems see the works of Kac, Kiefer & Wolfowitz [8], Durbin [4] or Martynov [12], [13].

A similar problem exists for the continuous time stochastic processes, which are widely used as mathematic models in many fields. The goodness of fit tests (GoF) are studied by many authors. For example Kutoyants [9] discusses some possibilities of the construction of such tests. In particular, he considers the Kolmogorov-Smirnov statistics and the Cramer-von Mises Statistics based on the continuous observation. Note that the Kolmogorov-Smirnov statistics for ergodic diffusion process was studied in Fournie [6] and in Fournie and Kutoyants [7]. However, due to the structure of the covariance of the limit process, the Kolmogorov-Smirnov statistics is not asymptotically distribution free in diffusion process models. More recently Kutoyants [10] has proposed a modification of the Kolmogorov-Smirnov statistics for diffusion models that became asymptotically distribution free. See also Dachian and Kutoyants [2] where they propose some GoF tests for diffusion and inhomogeneous Poisson processes with simple basic hypothesis. It was shown that these tests are asymptotically distribution free. In the case of Ornstein-Uhlenbeck process Kutoyants showed that the Cramer-von Mizes type tests are asymptotically parameter free [11]. Another test was studied by Negri and Nishiyama [15].

2 Main Results

Suppose that we observe an ergodic diffusion process, solution to the following stochastic differential equation

$$\underset{\cdot}{\mathrm{X}}_t = S(X_t)\underset{\cdot}{\mathrm{t}} + \underset{\cdot}{\mathrm{W}}_t, \quad X_0,\ 0 \le t \le T. \tag{2.1}$$

We want to test the following null hypothesis

$$\mathcal{H}_0 \quad : \qquad S(x) = S_*(x - \vartheta), \quad \vartheta \in \Theta,$$

where $S_*(\cdot)$ is some known function and the shift parameter ϑ is unknown. We suppose that $0 \in \Theta = (\alpha, \beta)$. Let us introduce the family

$$\mathcal{S}(\Theta) = \{S_*(x - \vartheta), \quad \vartheta \in \Theta = (\alpha, \beta)\}.$$

The alternative is defined as

$$\mathcal{H}_1 \quad : \qquad S(\cdot) \notin \overline{\mathcal{S}(\Theta)},$$

where $\overline{\mathcal{S}(\Theta)} = \{S(x - \vartheta), \vartheta \in [\alpha, \beta]\}$.

We suppose that the trend coefficients $S(\cdot)$ of the observed diffusion process under both hypotheses satisfy the conditions:

$\mathcal{ES}$. *The function $S(\cdot)$ is locally bounded and for some $C > 0$,*

$$xS(x) \le C(1 + x^2).$$

and

$\mathcal{A}_0$. *The function $S(\cdot)$ satisfies*

$$\varlimsup_{|x| \to \infty} \operatorname{sgn}(x) S(x) < 0. \tag{2.2}$$

Remind that under the condition $\mathcal{ES}$, the equation (2.1) has a unique weak solution (See [5]). Moreover under the condition $\mathcal{A}_0$, the diffusion process is recurrent and its invariant density $f(x, \vartheta)$ under hypothesis $\mathcal{H}_0$ can be given explicitly (See [9], Theorem 1.16):

$$f(x, \vartheta) = \frac{1}{G(\vartheta)} \exp\left\{2 \int_\vartheta^x S_*(y - \vartheta)\underset{\cdot}{\mathrm{y}}\right\}.$$

Denote by ξ_ϑ a random variable (r.v.) having this density and the corresponding mathematic expectation by $\mathbf{E}_\vartheta$. To simplify the notations, for the case $\vartheta = 0$, we denote the density function as $f(x) = f(x, 0)$, and the

corresponding distribution function as $F(x)$; correspondingly the r.v. is ξ_0, and the mathematical expectation is $\mathbf{E}_0$. Denote $\mathcal{P}$ as the class of functions having polynomial majorants i.e.

$$\mathcal{P} = \{h(\cdot) : \ |h(x)| \leq C(1 + |x|^p)\},$$

with some $p > 0$. Let $h'(x)$ the derivative of $h(x)$ w.r.t. x.

Let us fix some $\varepsilon \in (0, 1)$, and denote by $\mathcal{K}_\varepsilon$ the class of tests ψ_T of asymptotic size ε, i.e.

$$\mathbf{E}_0 \psi_T = \varepsilon + o(1).$$

Our object is to construct this kind of tests.

To verify the hypothesis $\mathcal{H}_0$, we propose two tests. The first one is based on the local time estimator (LTE) $\hat{f}_T(x)$ of the invariant density, which can be written as

$$\hat{f}_T(x) = \frac{1}{T}(|X_T - x| - |X_0 - x|) - \frac{1}{T}\int_0^T \mathrm{sgn}(X_t - x)\underset{\cdot}{\mathrm{X}}_t.$$

The unknown parameter is estimated via the maximum likelihood estimator (MLE) $\hat{\vartheta}_T$, which is defined as the solution of the equation

$$L(\hat{\vartheta}_T, X^T) = \sup_{\theta \in \Theta} L(\theta, X^T),$$

where $L(\vartheta, X^T)$ is the log-likelihood ratio

$$L(\vartheta, X^T) = \int_0^T S_*(X_t - \vartheta)\underset{\cdot}{\mathrm{X}}_t - \frac{1}{2}\int_0^T S_*(X_t - \vartheta)^2 \underset{\cdot}{\mathrm{t}}.$$

We give the following regularity conditions $\mathcal{A}$ to have the consistency and the asymptotical normality of the MLE:

Condition $\mathcal{A}$.

$\mathcal{A}_1$. *The function $S_*(\cdot)$ is continuously differentiable, the derivative $S_*'(\cdot) \in \mathcal{P}$ and is uniformly continuous in the following sense:*

$$\lim_{\nu \to 0} \sup_{|\tau| < \nu} \mathbf{E}_0 \left| S_*'(\xi_0) - S_*'(\xi_0 + \tau) \right|^2 = 0.$$

$\mathcal{A}_2$. *The Fisher information*

$$I = \mathbf{E}_0 S_*'(\xi_0)^2 > 0. \tag{2.3}$$

Moreover, for any $\nu > 0$

$$\inf_{|\tau|>\nu} \mathbf{E}_0 \big(S_*(\xi_0) - S_*(\xi_0 + \tau)\big)^2 > 0.$$

Denote the statistic based on the LTE as follows

$$\delta_T = T \int_{-\infty}^{\infty} \left(\hat{f}_T(x) - f(x - \hat{\vartheta}_T)\right)^2 \underset{\cdot}{\mathrm{x}},$$

we will prove that under hypothesis $\mathcal{H}_0$, it converges in distribution to

$$\delta = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \left(2f(x)\frac{\mathbb{1}_{\{y>x\}} - F(y)}{\sqrt{f(y)}} - \frac{1}{I} S_*'(y)\sqrt{f(y)}f'(x)\right) \underset{\cdot}{\mathrm{W}}(y)\right)^2 \underset{\cdot}{\mathrm{x}}, \tag{2.4}$$

with $W(y) = W_1(y),\ y \in \mathbb{R}^+,\ W(y) = W_2(-y),\ y \in \mathbb{R}^-$, where W_1 and W_2 are independent Wiener processes. The Cramer-von Mises type test is defined as

$$\psi_T = \mathbb{1}_{\{\delta_T > d_\varepsilon\}},$$

where d_ε is the $1 - \varepsilon$ quantile of the distribution of δ, that is the solution of the following equation

$$\mathbf{P}\Big(\delta \geq d_\varepsilon\Big) = \varepsilon. \tag{2.5}$$

The main result for the Cramer von Mises test based on local time estimator is the following:

Theorem 2.1. *Let the conditions $\mathcal{ES}$, $\mathcal{A}_0$ and $\mathcal{A}$ be fulfilled, then the test $\psi_T = \mathbb{1}_{\{\delta_T > d_\varepsilon\}}$ belongs to $\mathcal{K}_\varepsilon$.*

The theorem is proved in Section 3.

Note that neither δ nor d_ε depends on the unknown parameter. This allows us to conclude that the test is APF.

The second test is based on the same MLE and the empirical distribution function (EDF):

$$\hat{F}_T(x) = \frac{1}{T}\int_0^T \mathbb{1}_{\{X_t < x\}} \underset{\cdot}{\mathrm{t}}.$$

The corresponding statistic is

$$\Delta_T = T \int_{-\infty}^{\infty} \left(\hat{F}_T(x) - F(x - \hat{\vartheta}_T)\right)^2 \underset{\cdot}{\mathrm{x}},$$

which converges in distribution to

$$\Delta = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\left(2\frac{F(y\wedge x)-F(y)F(x)}{\sqrt{f(y)}}-\frac{1}{I}S_*'(y)\sqrt{f(y)}f(x)\right)\underset{\cdot}{\mathrm{W}}(y)\right)^2\underset{\cdot}{\mathrm{x}}. \tag{2.6}$$

Thus we propose the Cramer-von Mises type test

$$\Psi_T = 1\!\!1_{\{\Delta_T > c_\varepsilon\}},$$

where c_ε is the solution of the equation

$$\mathbf{P}\Big(\Delta \geq c_\varepsilon\Big) = \varepsilon. \tag{2.7}$$

The main result for the Cramer von Mises test based on empirical distribution function estimator is the following:

Theorem 2.2. *Under conditions $\mathcal{ES}$, $\mathcal{A}_0$ and $\mathcal{A}$, the test $\Psi_T = 1\!\!1_{\{\Delta_T > c_\varepsilon\}}$ belongs to $\mathcal{K}_\varepsilon$.*

The theorem is proved In Section 4.

3 Proof of Theorem 2.1

In this section, we study the test $\psi_T = 1\!\!1_{\{\delta_T > d_\varepsilon\}}$, where

$$\delta_T = T\int_{-\infty}^{\infty}\left(\hat{f}_T(x) - f(x-\hat{\vartheta}_T)\right)^2\underset{\cdot}{\mathrm{x}}.$$

Under the basic hypothesis $\mathcal{H}_0$, the density of the invariant law can be presented as follows:

$$\begin{aligned} f(x,\vartheta) &= \frac{\exp\{2\int_\vartheta^x S_*(y-\vartheta)\underset{\cdot}{\mathrm{y}}\}}{\int_{-\infty}^{\infty}\exp\{2\int_\vartheta^y S_*(z-\vartheta)\underset{\cdot}{\mathrm{z}}\}\underset{\cdot}{\mathrm{y}}} \\ &= \frac{\exp\{2\int_0^{x-\vartheta} S_*(y)\underset{\cdot}{\mathrm{y}}\}}{\int_{-\infty}^{\infty}\exp\{2\int_0^{y-\vartheta} S_*(z)\underset{\cdot}{\mathrm{z}}\}\underset{\cdot}{\mathrm{y}}} \\ &= f(x-\vartheta). \end{aligned}$$

Note that the distribution function of the process satisfies

$$F(x,\vartheta) = \int_{-\infty}^{x} f(y-\vartheta)\underset{\cdot}{\mathrm{y}} = \int_{-\infty}^{x-\vartheta} f(y)\underset{\cdot}{\mathrm{y}} = F(x-\vartheta).$$

In addition, for any integrable function h,

$$
\begin{aligned}
\mathbf{E}_\vartheta h(\xi_\vartheta - \vartheta) &= \int_{-\infty}^{\infty} h(x-\vartheta) f(x-\vartheta)\mathrm{d}x \\
&= \int_{-\infty}^{\infty} h(x) f(x)\mathrm{d}x = \mathbf{E}_0 h(\xi_0). \qquad (3.1)
\end{aligned}
$$

Note that the Fisher information in our case does not depend on the unknown parameter ϑ:

$$
I = \mathbf{E}_{\vartheta_0} S'_*(\xi_{\vartheta_0} - \vartheta_0)^2 = \mathbf{E}_0 S'_*(\xi_0)^2 > 0.
$$

where ϑ_0 is the true value of the unknown parameter.

From the condition $\mathcal{A}_0$, it follows that there exist some constants $A > 0$ and $\gamma > 0$ such that for all $|x| > A$,

$$
\operatorname{sgn}(x) S_*(x) < -\gamma. \qquad (3.2)
$$

It can be shown that for $x > A$,

$$
f(x) \;=\; \frac{1}{G(S_*)} \exp\left\{ 2 \left(\int_0^A + \int_A^x \right) S_*(y)\mathrm{d}y \right\} < C\mathrm{e}^{-2\gamma x}.
$$

Similar result can be deduced for $x < -A$, so we have

$$
f(x) < C\mathrm{e}^{-2\gamma|x|}, \quad \text{for } |x| > A. \qquad (3.3)
$$

Let the conditions $\mathcal{A}_0$ and $\mathcal{A}$ be fulfilled, then the MLE $\hat\vartheta_T$ is consistent, i.e., for any $\nu > 0$,

$$
\lim_{T\to\infty} \mathbf{P}_{\vartheta_0}\big\{ |\hat\vartheta_T - \vartheta_0| > \nu \big\} = 0;
$$

it is asymptotically normal

$$
\mathcal{L}_{\vartheta_0}\big\{ \sqrt{T}(\hat\vartheta_T - \vartheta_0) \big\} \Longrightarrow \mathcal{N}(0, I^{-1}); \qquad (3.4)
$$

and the moments converge i.e., for $p > 0$

$$
\lim_{T\to\infty} \mathbf{E}_{\vartheta_0} \left| \sqrt{T}(\hat\vartheta_T - \vartheta_0) \right|^p = \mathbf{E}_0 \left| \hat u \right|^p,
$$

where $\hat u \sim \mathcal{N}(0, I^{-1})$. The proof can be found in [9],Theorem 2.8. We can define

$$
\hat u = \frac{1}{I} \int_{-\infty}^{\infty} S'_*(y) \sqrt{f(y)}\mathrm{d}W(y),
$$

and denoted $\hat{u}_T = \sqrt{T}(\hat{\vartheta}_T - \vartheta_0)$, the asymptotical normality (3.4) can be written as

$$\mathcal{L}_{\vartheta_0}\left\{\hat{u}_T\right\} \Longrightarrow \mathcal{L}\left\{\hat{u}\right\}. \tag{3.5}$$

We define $\eta_T(x) = \sqrt{T}\left(\hat{f}_T(x) - f(x-\vartheta_0)\right)$. In [9] Theorem 4.11, we can find the following representation

$$\begin{aligned}
\eta_T(x) &= \sqrt{T}(\hat{f}_T(x) - f(x-\vartheta_0)) \\
&= 2\frac{f(x-\vartheta_0)}{\sqrt{T}}\int_{X_0}^{X_T}\left(\frac{1\!\!\mathrm{I}_{\{y>x\}} - F(y-\vartheta_0)}{f(y-\vartheta_0)}\right)\underset{\cdot}{\mathrm{y}} \\
&\quad -2\frac{f(x-\vartheta_0)}{\sqrt{T}}\int_0^T\left(\frac{1\!\!\mathrm{I}_{\{X_t>x\}} - F(X_t-\vartheta_0)}{f(X_t-\vartheta_0)}\right)\underset{\cdot}{\mathrm{W}}_t.
\end{aligned} \tag{3.6}$$

Let us put

$$M(y,x) = 2f(x)\frac{1\!\!\mathrm{I}_{\{y>x\}} - F(y)}{f(y)}.$$

Then $\eta_T(x)$ can be written as

$$\begin{aligned}
\eta_T(x) &= \frac{1}{\sqrt{T}}\int_{X_0}^{X_T} M(y-\vartheta_0, x-\vartheta_0)\underset{\cdot}{\mathrm{y}} \\
&\quad -\frac{1}{\sqrt{T}}\int_0^T M(X_t-\vartheta_0, x-\vartheta_0)\underset{\cdot}{\mathrm{W}}_t.
\end{aligned} \tag{3.7}$$

We can state

Lemma 3.1. *Let the condition $\mathcal{A}_0$ be fulfilled, then*

$$\int_{-\infty}^{\infty}\mathbf{E}_0\left(\int_0^{\xi_0} M(y,x)\underset{\cdot}{\mathrm{y}}\right)^2\underset{\cdot}{\mathrm{x}} < \infty.$$

Proof. Applying the estimate (3.3), for $x > A$,

$$\begin{aligned}
&\mathbf{E}_0\left(\int_0^{\xi_0} M(y,x)\underset{\cdot}{\mathrm{y}}\right)^2 \\
&= 4f(x)^2\int_{-\infty}^{\infty}\left(\int_0^z \frac{1\!\!\mathrm{I}_{\{y>x\}} - F(y)}{f(y)}\underset{\cdot}{\mathrm{y}}\right)^2 f(z)\underset{\cdot}{\mathrm{z}} \\
&= 4f(x)^2\left(\int_{-\infty}^{-A} + \int_{-A}^{A} + \int_A^x\right)\left(\int_0^z \frac{-F(y)}{f(y)}\underset{\cdot}{\mathrm{y}}\right)^2 f(z)\underset{\cdot}{\mathrm{z}} \\
&\quad +4f(x)^2\int_x^{\infty}\left(\int_0^x \frac{-F(y)}{f(y)}\underset{\cdot}{\mathrm{y}} + \int_x^z \frac{1-F(y)}{f(y)}\underset{\cdot}{\mathrm{y}}\right)^2 f(z)\underset{\cdot}{\mathrm{z}}
\end{aligned}$$

Further,

$$
\begin{aligned}
&f(x)^2 \int_{-\infty}^{-A} \left(\int_0^z \frac{-F(y)}{f(y)} \mathrm{d}y \right)^2 f(z) \mathrm{d}z \\
&= f(x)^2 \int_{-\infty}^{-A} \left(\left(\int_z^{-A} + \int_{-A}^0 \right) \frac{F(y)}{f(y)} \mathrm{d}y \right)^2 f(z) \mathrm{d}z \\
&\le f(x)^2 \int_{-\infty}^{-A} \left(\int_z^{-A} \int_{-\infty}^{y} \frac{1}{G} \exp \left(-2 \int_u^y S_*(v) \mathrm{d}v \right) \mathrm{d}u \mathrm{d}y + C_1 \right)^2 f(z) \mathrm{d}z \\
&\le f(x)^2 \int_{-\infty}^{-A} \left(C_2 \int_z^{-A} \int_{-\infty}^{y} \mathrm{e}^{-2\gamma(y-u)} \mathrm{d}u \mathrm{d}y + C_1 \right)^2 f(z) \mathrm{d}z \\
&\le C f(x)^2 \int_{-\infty}^{-A} (1+z)^2 f(z) \mathrm{d}z \le C f(x)^2 \le C \mathrm{e}^{-4\gamma x},
\end{aligned}
$$

moreover

$$
\begin{aligned}
&f(x)^2 \int_A^x \left(\int_0^z \frac{-F(y)}{f(y)} \mathrm{d}y \right)^2 f(z) \mathrm{d}z \\
&\le \int_A^x \left(\left(\int_0^A + \int_A^z \right) \frac{f(x)}{f(y)} \mathrm{d}y \right)^2 f(z) \mathrm{d}z \\
&\le \int_A^x \left(C_1 f(x) + C_2 \int_A^z \mathrm{e}^{-2\gamma(x-y)} \mathrm{d}y \right)^2 f(z) \mathrm{d}z \\
&\le \int_A^x \left(C_1 \mathrm{e}^{-2\gamma x} + C_2' \mathrm{e}^{-2\gamma(x-z)} - C_2' \mathrm{e}^{-2\gamma(x-A)} \right)^2 \cdot C \mathrm{e}^{-2\gamma z} \mathrm{d}z \\
&\le \mathrm{e}^{-4\gamma x} \int_A^x \left(C_3 \mathrm{e}^{2\gamma z} + C_4 \mathrm{e}^{-2\gamma z} \right) \mathrm{d}z \le C \mathrm{e}^{-2\gamma x},
\end{aligned}
$$

and finally

$$
\begin{aligned}
&f(x)^2 \int_x^{\infty} \left(\int_x^z \frac{1-F(y)}{f(y)} \mathrm{d}y \right)^2 f(z) \mathrm{d}z \\
&\le C f(x)^2 \int_x^{\infty} \left(\int_x^z \int_y^{\infty} \mathrm{e}^{-2\gamma(u-y)} \mathrm{d}u \mathrm{d}y \right)^2 \mathrm{e}^{-2\gamma z} \mathrm{d}z \\
&\le C f(x)^2 \int_x^{\infty} (z-x)^2 \mathrm{e}^{-2\gamma z} \mathrm{d}z \\
&\le C f(x)^2 \int_0^{\infty} s^2 \mathrm{e}^{-2\gamma(s+x)} \mathrm{d}s \le C \mathrm{e}^{-6\gamma x}.
\end{aligned}
$$

Then we have

$$
\mathbf{E}_0 \left(\int_0^{\xi_0} M(y,x) \mathrm{d}y \right)^2 \le C \mathrm{e}^{-2\gamma|x|} \quad \text{for } x > A. \tag{3.8}
$$

Similar estimate can be obtained for $x < -A$, therefore the result holds for $|x| > A$. We obtain finally

$$\begin{aligned}
&\int_{-\infty}^{\infty} \mathbf{E}_0 \left(\int_0^{\xi_0} M(y,x) \underset{\cdot}{\mathrm{y}} \right)^2 \underset{\cdot}{\mathrm{x}} \\
&= \left(\int_{-\infty}^{-A} + \int_{-A}^{A} + \int_{A}^{\infty} \right) \mathbf{E}_0 \left(\int_0^{\xi_0} M(y,x) \underset{\cdot}{\mathrm{y}} \right)^2 \underset{\cdot}{\mathrm{x}} \\
&\leq C_1 \int_{-\infty}^{-A} \mathrm{e}^{2\gamma x} \underset{\cdot}{\mathrm{x}} + C_2 + C_3 \int_A^{\infty} \mathrm{e}^{-2\gamma x} \underset{\cdot}{\mathrm{x}} < \infty.
\end{aligned}$$

This result yields directly the conditions $\mathcal{O}$ of Theorem 4.11 in [9]:

$$\mathbf{E}_{\vartheta_0} M(\xi_{\vartheta_0} - \vartheta_0, x - \vartheta_0)^2 = \mathbf{E}_0 M(\xi_0, x - \vartheta_0)^2 < \infty,$$

and

$$\mathbf{E}_{\vartheta_0} \left(\int_0^{\xi_{\vartheta_0}} M(y - \vartheta_0, x - \vartheta_0) \underset{\cdot}{\mathrm{y}} \right)^2 < \infty.$$

So we can deduce the convergence and the asymptotical normality of $\eta_T(x)$. In fact under the condition $\mathcal{A}_0$, the LTE $\hat{f}_T(x)$ is consistent and asymptotically normal, that is

$$\eta_T(x) = \sqrt{T} \left(\hat{f}_T(x) - f(x - \vartheta_0) \right) \Longrightarrow \eta(x - \vartheta_0),$$

where $\eta(x) \sim \mathcal{N}(0, d(x)^2)$, and

$$d(x)^2 = 4 f(x)^2 \mathbf{E}_0 \left(\frac{1\!\!\mathrm{I}_{\{\xi_0 > x\}} - F(\xi_0)}{f(\xi_0)} \right)^2.$$

Moreover

$$\begin{aligned}
&\mathbf{E}_{\vartheta_0} \left(\eta_T(x) \eta_T(y) \right) \\
&\quad = 4 f(x - \vartheta_0) f(y - \vartheta_0) \mathbf{E}_0 \left(\frac{\left(1\!\!\mathrm{I}_{\{\xi_0 > x - \vartheta_0\}} - F(\xi_0)\right) \left(1\!\!\mathrm{I}_{\{\xi_0 > y - \vartheta_0\}} - F(\xi_0)\right)}{f(\xi_0)^2} \right).
\end{aligned}$$

We can define

$$\eta(x) = \int_{-\infty}^{\infty} M(y,x) \sqrt{f(y)} \underset{\cdot}{\mathrm{W}}(y).$$

The distribution of $\eta(x)$ is $\mathcal{N}(0, \mathbf{E}_0 M(\xi_0, x)^2)$, and we have the following convergence

$$\eta_T(x) \Longrightarrow \eta(x - \vartheta_0). \tag{3.9}$$

For $\hat{u}_T$ and $\eta_T(x)$, we need more than (3.5) and convergence (3.9).

Lemma 3.2. *Let conditions $\mathcal{A}_0$ and $\mathcal{A}$ be fulfilled, then $(\eta_T(x_1), ..., \eta_T(x_k), \hat{u}_T)$ is asymptotically normal:*

$$\mathcal{L}\left(\eta_T(x_1), ..., \eta_T(x_k), \hat{u}_T\right) \Longrightarrow \mathcal{L}\left(\eta(x_1-\vartheta_0), ..., \eta(x_k-\vartheta_0), \hat{u}\right),$$

for any $\mathbf{x} = \{x_1, x_2, ..., x_k\} \in \mathbb{R}^k$.

Proof. The first integral in (3.7) converges to zero, so we only need to verify the convergence for the part of Itô integral. Let us denote for simplicity

$$\eta_T^0(x) = \frac{1}{\sqrt{T}} \int_0^T M(X_t - \vartheta_0, x) \underset{\cdot}{\mathrm{W}}_t.$$

It is sufficient to verify that for any $\mathbf{x} = \{x_1, x_2, ..., x_k\}$,

$$\left(\eta_T^0(x_1), ..., \eta_T^0(x_k), \hat{u}_T\right) \Longrightarrow \left(\eta(x_1), ..., \eta(x_k), \hat{u}\right). \tag{3.10}$$

Remember that $\hat{u}_T$ can be defined as follows,

$$Z_T(\hat{u}_T) = \sup_{u \in \mathbb{U}_T} Z_T(u), \quad \mathbb{U}_T = \{u : \vartheta + \frac{u}{\sqrt{T}} \in \Theta\}, \tag{3.11}$$

where

$$Z_T(u) = \frac{\underset{\cdot}{\mathbf{P}}^T_{\vartheta + \frac{u}{\sqrt{T}}}}{\underset{\cdot}{\mathbf{P}}^T_{\vartheta}}(X^T) = \exp\left\{u\Lambda_T - \frac{u^2}{2} I + r_T\right\}.$$

Here $\Lambda_T = \frac{1}{\sqrt{T}} \int_0^T S_*'(X_t - \vartheta_0) \underset{\cdot}{\mathrm{W}}_t$ and $r_T \longrightarrow 0$. It was proved in [9], Theorem 2.8 that $Z_T(\cdot)$ converges in distribution to $Z(\cdot)$, where

$$Z(u) = \exp\left\{u\Lambda - \frac{u^2}{2} I\right\},$$

where Λ is a r.v. with normal distribution $\mathcal{N}(0, I)$, which can be written as

$$\Lambda = \int_{-\infty}^{\infty} S_*'(y) \sqrt{f(y)} \underset{\cdot}{\mathrm{W}}(y).$$

Therefore

$$\hat{u}_T \Longrightarrow \hat{u} = \frac{\Lambda}{I}.$$

Take $\mathbf{u} = \{u_1, u_2, ..., u_m\}$. We have to verify that the joint finite-dimensional distribution of Y_T

$$Y_T = \left(\eta_T^0(x_1), \eta_T^0(x_2), ..., \eta_T^0(x_k), Z_T(u_1), Z_T(u_2), ..., Z_T(u_m)\right)$$

converges to the finite-dimensional distribution of Y

$$Y = (\eta(x_1), \eta(x_2), ..., \eta(x_k), Z(u_1), Z(u_2), ..., Z(u_m)) \, .$$

Note that the only stochastic term in $Z_T(u)$ is Λ_T, so (3.10) is equivalent to

$$\left(\eta_T^0(x_1), \eta_T^0(x_2), ..., \eta_T^0(x_k), \Lambda_T\right) \Longrightarrow (\eta(x_1), \eta(x_2), ..., \eta(x_k), \Lambda) \, . \qquad (3.12)$$

Take $\lambda = \{\lambda_1, \lambda_2, ..., \lambda_{k+1}\}$, and put

$$h(y, \mathbf{x}, \lambda) = \sum_{l=1}^{k} \lambda_l M(y, x_l) + \lambda_{k+1} S_*'(y).$$

We have

$$\begin{aligned}
&\mathbf{E}_{\vartheta_0} h(\xi_{\vartheta_0} - \vartheta_0, \mathbf{x}, \lambda)^2 = \mathbf{E}_0 h(\xi_0, \mathbf{x}, \lambda)^2 \\
&= \int_{-\infty}^{\infty} \left(\sum_{l=1}^{k} \lambda_l M(y, x_l) + \lambda_{k+1} S_*'(y) \right)^2 f(y) \underset{\cdot}{\mathrm{y}} \\
&= \int_{-\infty}^{\infty} \left(\sum_{l=1}^{k} 2\lambda_l f(x_l) \frac{1\!\!\mathrm{I}_{\{y > x_l\}} - F(y)}{\sqrt{f(y)}} + \lambda_{k+1} S_*'(y) \sqrt{f(y)} \right)^2 f(y) \underset{\cdot}{\mathrm{y}} \\
&= \int_{-\infty}^{\infty} \left(\sum_{l=1}^{k} \sum_{m=1}^{k} 4\lambda_l \lambda_m f(x_l) f(x_m) \frac{(1\!\!\mathrm{I}_{\{y > x_l\}} - F(y))(1\!\!\mathrm{I}_{\{y > x_m\}} - F(y))}{f(y)} \right. \\
&\quad \left. + \sum_{l=1}^{k} \lambda_l \lambda_{k+1} \left(1\!\!\mathrm{I}_{\{y > x_l\}} - F(y)\right) S_*'(y) + \lambda_{k+1}^2 S_*'(y)^2 f(y) \right) \underset{\cdot}{\mathrm{y}} < \infty.
\end{aligned}$$

The law of large number gives us

$$\frac{1}{T} \int_0^T h(X_t - \vartheta_0, \mathbf{x}, \lambda)^2 \underset{\cdot}{\mathrm{t}} \longrightarrow \mathbf{E}_0 h(\xi_0, \mathbf{x}, \lambda)^2.$$

Moreover, the central limit theorem for stochastic integral gives us

$$\frac{1}{\sqrt{T}} \int_0^T h(X_t - \vartheta_0, \mathbf{x}, \lambda) \underset{\cdot}{\mathrm{W}}_t \Longrightarrow \mathcal{N}\left(0, \mathbf{E}_0 h(\xi_0, \mathbf{x}, \lambda)^2\right).$$

In addition $\sum\limits_{l=1}^{k} \lambda_l \eta(x_l) + \lambda_{k+1} \Lambda$ is a zero mean normal r.v. with variance

$$\begin{aligned}
&\mathbf{E}_0 \left(\sum_{l=1}^{k} \lambda_l \eta(x_l) + \lambda_{k+1} \Lambda \right)^2 \\
&= \sum_{l=1}^{k} \sum_{m=1}^{k} \lambda_l \lambda_m \mathbf{E}_0 \left(\eta(x_l) \eta(x_m)\right) + \sum_{l=1}^{k} \lambda_l \lambda_{k+1} \mathbf{E}_0(\eta(x_l) \Lambda) + \lambda_{k+1}^2 \mathbf{E}_0(\Lambda)^2.
\end{aligned}$$

Furthermore

$$
\begin{aligned}
&\mathbf{E}_0\left(\eta(x_l)\eta(x_m)\right) \\
&\quad = 4f(x_l)f(x_l)\int_{-\infty}^{\infty}\frac{(1\!\!\mathrm{I}_{\{y>x_l\}}-F(y))(1\!\!\mathrm{I}_{\{y>x_m\}}-F(y))}{f(y)}\underset{\cdot}{\mathrm{y}},
\end{aligned}
$$

and

$$
\mathbf{E}_0(\eta(x_l)\Lambda) = -2f(x_l)\int_{-\infty}^{\infty}(1\!\!\mathrm{I}_{\{y>x_l\}}-F(y))S_*'(y)\underset{\cdot}{\mathrm{y}},
$$

$$
\mathbf{E}_0(\Lambda)^2 = \int_{-\infty}^{\infty}S_*'(y)^2 f(y)\underset{\cdot}{\mathrm{y}}.
$$

We find that

$$
\mathbf{E}_{\vartheta_0}h(\xi_{\vartheta_0}-\vartheta_0,\mathbf{x},\lambda)^2 = \mathbf{E}_0 h(\xi_0,\mathbf{x},\lambda)^2 = \mathbf{E}_0\left(\sum_{l=1}^{k}\lambda_l\eta(x_l)+\lambda_{k+1}\Lambda\right)^2 .
$$

This is as to say

$$
\sum_{l=1}^{k}\lambda_l\eta_T^0(x_l)+\lambda_{k+1}\Lambda_T \Longrightarrow \sum_{l=1}^{k}\lambda_l\eta(x_l)+\lambda_{k+1}\Lambda
$$

thus (3.10) follows from this last convergence in distribution, and so the lemma is proved.

Lemma 3.3. *Let conditions $\mathcal{A}_0$ and $\mathcal{A}$ be fulfilled, then*

$$
\mathcal{L}\left\{\int_{-\infty}^{\infty}\left(\eta_T^0(x)-\hat{u}_T f'(x)\right)^2\underset{\cdot}{\mathrm{x}}\right\} \Longrightarrow \mathcal{L}\left\{\int_{-\infty}^{\infty}\left(\eta(x)-\hat{u}f'(x)\right)^2\underset{\cdot}{\mathrm{x}}\right\}
$$

Proof. Denote $\zeta_T(x)=\eta_T^0(x)-\hat{u}_T f'(x)$ and $\zeta(x)=\eta(x)-\hat{u}f'(x)$, we will prove the following properties

i) For $x,y\in[-L,L]$ and $|x-y|\le 1$,

$$
\mathbf{E}_{\vartheta_0}|\zeta_T(x)^2-\zeta_T(y)^2|^2 \le C|x-y|^{\delta}, \qquad \text{with some } \delta>0. \tag{3.13}
$$

ii) $\forall\varepsilon>0,\ \exists L>0$, such that

$$
\mathbf{E}_{\vartheta_0}\int_{\{|x|>L\}}\zeta_T(x)^2\underset{\cdot}{\mathrm{x}} < \varepsilon, \quad \forall T>0. \tag{3.14}
$$

From i) it follows the convergence in every bounded set $[-L, L]$:

$$\mathcal{L}\Big\{ \int_{-L}^{L} \zeta_T(x)^2 \underset{\cdot}{\mathrm{x}} \Big\} \Longrightarrow \mathcal{L}\Big\{ \int_{-L}^{L} \zeta(x)^2 \underset{\cdot}{\mathrm{x}} \Big\}.$$

The result in i) along with ii) gives us the result.

First we prove i). We have

$$\mathbf{E}_{\vartheta_0} \left(\zeta_T(x)^2 \right) \leq 2\mathbf{E}_{\vartheta_0} \eta_T^0(x)^2 + 2f(x)^2 \mathbf{E}_{\vartheta_0} \hat{u}_T^2 \leq C.$$

$$\begin{aligned}
&\mathbf{E}_{\vartheta_0} \left| \zeta_T(x)^2 - \zeta_T(y)^2 \right|^2 \\
&\quad = \mathbf{E}_{\vartheta_0} \left(|\zeta_T(x) + \zeta_T(y)|^2 |\zeta_T(x) - \zeta_T(y)|^2 \right) \\
&\quad \leq C\mathbf{E}_{\vartheta_0} |\zeta_T(x) - \zeta_T(y)|^2 \\
&\quad \leq C(f'(x) - f'(y))^2 \mathbf{E}_{\vartheta_0} |\hat{u}_T|^2 + \mathbf{E}_{\vartheta_0} |(\eta_T^0(x) - \eta_T^0(y))|^2.
\end{aligned}$$

For the first part, let us recall the following result, given in [9], page 119: for any $p > 0,\ R > 0$, chosen N sufficiently large, we have

$$\mathbf{P}_{\vartheta_0}^T \left\{ |\hat{u}_T|^p > R \right\} \leq \frac{C_N}{R^{N/p}}.$$

Now, denoted $F_T(u)$ the distribution of $|\hat{u}_T|$, we have

$$\begin{aligned}
\mathbf{E}_{\vartheta_0} |\hat{u}_T|^p &= \int_0^\infty u^p \underset{\cdot}{\mathrm{F}}_T(u) \leq 1 - \int_1^\infty u^p \underset{\cdot}{[}1 - F_T(u)] \\
&\leq 1 - [1 - F_T(1)] + p \int_1^\infty u^{p-1} \frac{C_N}{u^{N/p}} \underset{\cdot}{\mathrm{u}} \leq C. \qquad (3.15)
\end{aligned}$$

Remember that under condition $\mathcal{A}_1$, S_* and f are sufficiently smooth. So, for $x, y \in [-L, L]$ we can write

$$|f(x) - f(y)| \ = \ |f'(z)(x-y)| = |2S_*(z)f(z)(x-y)| \leq C|x-y|,$$

and

$$|f'(x) - f'(y)| = |f''(z)(x-y)| = \left| 4f(z)S_*^2(z) + 2f(z)S_*'(z) \right| |x-y| \leq C|x-y|.$$

So we have

$$(f'(x) - f'(y))^2 \mathbf{E}_{\vartheta_0} |\hat{u}_T|^2 \leq C|x-y|^2.$$

For the second part, we can write

$$\begin{aligned}
&\mathbf{E}_{\vartheta_0}|(\eta_T^0(x)-\eta_T^0(y))|^2\\
&\quad = C_1\mathbf{E}_{\vartheta_0}\left(\frac{1}{\sqrt{T}}\int_0^T (M(X_t-\vartheta_0,x)-M(X_t-\vartheta_0,y))\underset{\cdot}{\mathrm{W}}_t\right)^2\\
&\quad \le \frac{C_1}{T}\int_0^T \mathbf{E}_{\vartheta_0}\left(M(X_t-\vartheta_0,x)-M(X_t-\vartheta_0,y)\right)^2 \underset{\cdot}{\mathrm{t}}\\
&\quad = C_1\mathbf{E}_0\left(M(\xi_0,x)-M(\xi_0,y)\right)^2.
\end{aligned}$$

Suppose that $x\le y$,

$$\begin{aligned}
&\mathbf{E}_0\left(M(\xi_0,x)-M(\xi_0,y)\right)^2\\
&\quad = \int_{-\infty}^x \left(2\frac{F(z)}{f(z)}(f(x)-f(y))\right)^2 f(z)\underset{\cdot}{\mathrm{z}}\\
&\qquad + \int_x^y \left(2\frac{1}{f(z)}\left((1-F(z))f(x)+F(z)f(y)\right)\right)^4 f(z)\underset{\cdot}{\mathrm{z}}\\
&\qquad + \int_y^\infty \left(2\frac{1-F(z)}{f(z)}(f(x)-f(y))\right)^2 f(z)\underset{\cdot}{\mathrm{z}}\\
&\quad \le C_1(x-y)^4+C_2(x-y)+C_3(x-y)^2\le C(y-x).
\end{aligned}$$

Similar result holds for $x>y$. Then we obtain

$$\mathbf{E}_{\vartheta_0}\left|\eta_T^0(x)^2-\eta_T^0(y)^2\right|^2\le C|x-y|,\quad x,y\in\mathbb{R}.$$

Thus we have

$$\mathbf{E}_{\vartheta_0}\left|\zeta_T(x)^2-\zeta_T(y)^2\right|^2\le C|x-y|.$$

Now we prove ii). As in Lemma 3.1, we can deduce that

$$\mathbf{E}_0 M(\xi_0,x)^2\le C\mathrm{e}^{-2\gamma x},\quad \text{for } x>A.$$

So for $L>A$,

$$\begin{aligned}
\mathbf{E}_{\vartheta_0}\int_L^\infty \left(\eta_T^0(x)\right)^2\underset{\cdot}{\mathrm{x}} &= \mathbf{E}_{\vartheta_0}\int_L^\infty\left(\frac{1}{\sqrt{T}}\int_0^T M(X_t-\vartheta_0,x)\underset{\cdot}{\mathrm{W}}_t\right)^2\underset{\cdot}{\mathrm{x}}\\
&\le C\int_L^\infty \mathbf{E}_0 M(\xi_0,x)^2\underset{\cdot}{\mathrm{x}}\le C\int_L^\infty \mathrm{e}^{-2\gamma x}\underset{\cdot}{\mathrm{x}}\le C\mathrm{e}^{-2\gamma L}.
\end{aligned}$$

Note that $f'(x)=2S_*(x)f(x)$ and along with (3.15) we get

$$\begin{aligned}
&\int_L^\infty \mathbf{E}_{\vartheta_0}\left(\eta_T^0(x)-f'(x)\hat{u}_T\right)^2\underset{\cdot}{\mathrm{x}}\\
&\quad\le \int_L^\infty\left(2\mathbf{E}_{\vartheta_0}\eta_T(x)^2+2f'(x)\mathbf{E}_{\vartheta_0}\hat{u}_T^2\right)\underset{\cdot}{\mathrm{x}}\\
&\quad\le \int_L^\infty C\mathrm{e}^{-2\gamma x}\underset{\cdot}{\mathrm{x}} = C\mathrm{e}^{-2\gamma L}.
\end{aligned}$$

For any $\varepsilon > 0$, take $L = -\frac{\ln(\varepsilon/C)}{2\gamma} \vee A$, then we have (3.14).

Proof of Theorem 2.1.

We can write

$$
\begin{aligned}
\delta_T &= T\int_{-\infty}^{\infty} (\hat{f}_T(x) - f(x-\hat{\vartheta}_T))^2 \underset{\cdot}{\mathrm{x}} \\
&= T\int_{-\infty}^{\infty} \Big((\hat{f}_T(x) - f(x-\vartheta_0)) + (f(x-\vartheta_0) - f(x-\hat{\vartheta}_T))\Big)^2 \underset{\cdot}{\mathrm{x}} \\
&= \int_{-\infty}^{\infty} \Big(\sqrt{T}(\hat{f}_T(x) - f(x-\vartheta_0)) - \sqrt{T}(\hat{\vartheta}_T - \vartheta_0) f'(x-\tilde{\vartheta}_T)\Big)^2 \underset{\cdot}{\mathrm{x}} \\
&= \int_{-\infty}^{\infty} \Big(\eta_T(x) - \hat{u}_T f'(x-\tilde{\vartheta}_T)\Big)^2 \underset{\cdot}{\mathrm{x}}.
\end{aligned}
$$

See that

$$
\begin{aligned}
&\mathbf{E}_{\vartheta_0} \int_{-\infty}^{\infty} \Big(\hat{u}_T^2 |f'(x-\tilde{\vartheta}_T) - f'(x-\vartheta_0)|^2\Big) \underset{\cdot}{\mathrm{x}} \\
&\quad = \mathbf{E}_{\vartheta_0} \int_{-\infty}^{\infty} \Big(\hat{u}_T^2 f''(x-\vartheta_T^*)^2 (\tilde{\vartheta}_T - \vartheta_0)^2\Big) \underset{\cdot}{\mathrm{x}},
\end{aligned}
$$

and that $f'(x-\vartheta) = S_*(x-\vartheta) f(x-\vartheta)$, $f''(x,\vartheta) = S_*'(x-\vartheta) f(x-\vartheta) + S_*(x-\vartheta)^2 f(x-\vartheta)$, the smoothness of $S_*(\cdot)$ gives us the convergence

$$
\mathbf{E}_{\vartheta_0} \int_{-\infty}^{\infty} \Big(\hat{u}_T^2 |f'(x-\tilde{\vartheta}_T) - f'(x-\vartheta_0)|^2\Big) \underset{\cdot}{\mathrm{x}} \longrightarrow 0.
$$

Applying Lemma 3.1 and Lemma 3.3 we get

$$
\begin{aligned}
\delta_T &= \int_{-\infty}^{\infty} \left(\eta_T^0(x-\vartheta_0) - \hat{u}_T f'(x-\vartheta_0)\right)^2 \underset{\cdot}{\mathrm{x}} + o(1) \\
&\Longrightarrow \int_{-\infty}^{\infty} \left(\eta(x-\vartheta_0) - \hat{u} f'(x-\vartheta_0)\right)^2 \underset{\cdot}{\mathrm{x}} \\
&= \int_{-\infty}^{\infty} \left(\eta(y) - \hat{u} f'(y)\right)^2 \underset{\cdot}{\mathrm{y}} = \delta.
\end{aligned}
$$

We see that the limit of the statistic δ does not depend on ϑ_0, and the test

$\psi_T = \mathbb{1}_{\{\delta_T \geq d_\varepsilon\}}$ with d_ε defined by

$$\mathbf{P}\Big(\delta \geq d_\varepsilon\Big) = \varepsilon$$

belongs to $\mathcal{K}_\varepsilon$.

The same procedure can be applied with other estimators of the unknown parameter and of the invariant density, provided that they are consistent and asymptotically normal. For example, we can take the minimum distance estimator (MDE) ϑ_T^* for ϑ_0:

$$\vartheta_T^* = \arg\inf_{\theta\in\Theta} \|\hat{F}(\cdot) - F(\theta,\cdot)\|,$$

and the kernel estimators $\bar{f}_T(x)$ as estimator for the invariant density

$$\bar{f}_T(x) = \frac{1}{\sqrt{T}} \int_0^T K(\sqrt{T}(X_t - x))\underset{\cdot}{t}.$$

Under some regularity conditions, the MDE $\hat{\vartheta}_T^*$ is asymptotically normal (See [7] or [9]):

$$u_T^* = \sqrt{T}(\vartheta_T^* - \vartheta_0) \Longrightarrow \hat{u}^* \sim \mathcal{N}(0, R(\vartheta_0)).$$

Also if we do not present explicitly $R(\cdot)$ here, it can be verified that $R(\vartheta) = R(0)$ does not depend on ϑ. The kernel estimator $\bar{f}_T(x)$ has the same asymptotic properties of the LTE (See [9]). Then we can construct the statistic

$$\mu_T = T\int_{-\infty}^{\infty} \left(\bar{f}(x) - f(x - \vartheta_T^*)\right)^2 \underset{\cdot}{x},$$

which converges to

$$\mu = \int_{-\infty}^{\infty} \left(\eta(x) - u^* f'(x)\right)^2 \underset{\cdot}{x},$$

that does not depend on the unknown parameter. So that the test $\mathbb{1}_{\{\mu_T > k_\varepsilon\}}$ with k_ε the solution of the equation

$$\mathbf{P}\left(\mu > k_\varepsilon\right) = \varepsilon$$

belongs to $\mathcal{K}_\varepsilon$.

4 Proof of Theorem 2.2

In this section, we study the GoF test $\Psi_T = \mathbb{1}_{\{\Delta_T \geq c_\varepsilon\}}$ defined by the statistic

$$\Delta_T = T \int_{-\infty}^{\infty} \left(\hat{F}_T(x) - F(x - \hat{\vartheta}_T)\right)^2 \underset{\cdot}{\mathrm{x}},$$

where $\hat{F}_T(x)$ is the empirical distribution function:

$$\hat{F}_T(x) = \frac{1}{T} \int_0^T \mathbb{1}_{\{X_t < x\}} \underset{\cdot}{\mathrm{t}}.$$

Denote $\eta_T^F(x) = \sqrt{T}(\hat{F}_T(x) - F(x - \vartheta_0))$ and

$$H(z, x) = 2\frac{F(z \wedge x) - F(z)F(x)}{f(z)}.$$

In [9] Theorem 4.6, the following equality is presented:

$$\begin{aligned} \eta_T^F(x) &= \frac{2}{\sqrt{T}} \int_{X_0}^{X_T} \frac{F((z \wedge x) - \vartheta_0) - F(z - \vartheta_0)F(x - \vartheta_0)}{f(z - \vartheta_0)} \underset{\cdot}{\mathrm{z}} \\ &\quad - \frac{2}{\sqrt{T}} \int_0^T \frac{F((X_t \wedge x) - \vartheta_0) - F(X_t - \vartheta_0)F(x - \vartheta_0)}{f(X_t - \vartheta_0)} \underset{\cdot}{\mathrm{W}}_t. \end{aligned}$$

Then

$$\begin{aligned} \eta_T^F(x) &= \frac{2}{\sqrt{T}} \int_{X_0}^{X_T} \frac{F((z - \vartheta_0) \wedge (x - \vartheta_0)) - F(z - \vartheta_0)F(x - \vartheta_0)}{f(z - \vartheta_0)} \underset{\cdot}{\mathrm{z}} \\ &\quad - \frac{2}{\sqrt{T}} \int_0^T \frac{F((X_t - \vartheta_0) \wedge (x - \vartheta_0)) - F(X_t - \vartheta_0)F(x - \vartheta_0)}{f(X_t - \vartheta_0)} \underset{\cdot}{\mathrm{W}}_t \\ &= \frac{1}{\sqrt{T}} \left(\int_0^{X_T} H(z - \vartheta_0, x - \vartheta_0) \underset{\cdot}{\mathrm{z}} - \int_0^{X_0} H(z - \vartheta_0, x - \vartheta_0) \underset{\cdot}{\mathrm{z}} \right) \\ &\quad - \frac{1}{\sqrt{T}} \int_0^T H(X_t - \vartheta_0, x - \vartheta_0) \underset{\cdot}{\mathrm{W}}_t. \end{aligned}$$

Using (3.2) we have, for $x > A$,

$$1 - F(x) = C \int_x^{\infty} \exp\left(2 \int_0^y S_*(r) \underset{\cdot}{\mathrm{r}}\right) \underset{\cdot}{\mathrm{y}} \leq C \mathrm{e}^{-2\gamma x},$$

and

$$\frac{1 - F(x)}{f(x)} \leq C \int_x^{\infty} \mathrm{e}^{-2\gamma(y - x)} \underset{\cdot}{\mathrm{y}} \leq C.$$

For $x < -A$ we have $F(x) \leq C\mathrm{e}^{-2\gamma|x|}$ and we can write

$$\frac{F(x)}{f(x)} = C \int_{-\infty}^{x} \exp(2 \int_{x}^{y} S_*(r)\underset{\cdot}{\mathrm{r}})\underset{\cdot}{\mathrm{y}} \leq C.$$

These inequalities allow us to deduce the following bounds

$$\mathbf{E}_{\vartheta_0} H(\xi_{\vartheta_0} - \vartheta_0, x)^2 = \mathbf{E}_0 H(\xi_0, x)^2 < \mathrm{e}^{-\gamma|x|}, \quad |x| > A. \tag{4.1}$$

and

$$\mathbf{E}_{\vartheta_0} \left(\int_0^{\xi_{\vartheta_0} - \vartheta_0} H(z,x)\underset{\cdot}{\mathrm{z}} \right)^2 = \mathbf{E}_0 \left(\int_0^{\xi_0} H(z,x)\underset{\cdot}{\mathrm{z}} \right)^2 \leq C\mathrm{e}^{-\gamma|x|}, \quad |x| > A. \tag{4.2}$$

Moreover

$$\int_{-\infty}^{\infty} \mathbf{E}_0 \left(\int_0^{\xi_0} H(z,x)\underset{\cdot}{\mathrm{z}} \right)^2 \underset{\cdot}{\mathrm{x}} \leq \infty. \tag{4.3}$$

Hence we get the asymptotic normality of $\eta_T^F(x)$:

$$\eta_T^F(x) \Longrightarrow \eta^F(x - \vartheta_0) \sim \mathcal{N}(0, 4\mathbf{E}_0 \left(H(\xi_0, x - \vartheta_0)\right)^2).$$

As in Lemma 3.2 and Lemma 3.3, if conditions $\mathcal{A}$ and $\mathcal{A}_0$ hold, we can show the convergence of the vector $(\eta_T^F(x_1), ..., \eta_T^F(x_k), \hat{u}_T)$:

$$\mathcal{L}_{\vartheta_0} \left(\eta_T^F(x_1), ..., \eta_T^F(x_k), \hat{u}_T \right) \Longrightarrow \mathcal{L}_{\vartheta_0} \left(\eta^F(x_1 - \vartheta_0), ..., \eta_T^F(x_k - \vartheta_0), \hat{u} \right)$$

and the convergence of the integral:

$$\mathcal{L}_{\vartheta_0} \{ \int_{-\infty}^{\infty} \left(\eta_T^F(x) - \hat{u}_T f(x - \vartheta_0) \right)^2 \underset{\cdot}{\mathrm{x}} \} \quad \Longrightarrow \quad \mathcal{L} \left\{ \int_{-\infty}^{\infty} \left(\eta^F(x) - \hat{u} f(x) \right)^2 \underset{\cdot}{\mathrm{x}} \right\}.$$

We obtain finally

$$\begin{aligned}
\Delta_T \quad &= \quad T \int_{-\infty}^{\infty} (\hat{F}_T(x) - F(x, \hat{\vartheta}_T))^2 \underset{\cdot}{\mathrm{x}} \\
&= \quad \int_{-\infty}^{\infty} \left[\sqrt{T}(\hat{F}_T(x) - F(x - \vartheta_0)) - \sqrt{T}(\hat{\vartheta}_T - \vartheta_0) F'(x - \tilde{\vartheta}_T) \right]^2 \underset{\cdot}{\mathrm{x}} \\
&= \quad \int_{-\infty}^{\infty} \left[\eta_T^F(x) - \hat{u}_T f(x - \tilde{\vartheta}_T) \right]^2 \underset{\cdot}{\mathrm{x}} \\
&= \quad \int_{-\infty}^{\infty} \left[\eta_T^F(x) - \hat{u}_T f(x - \vartheta_0) \right]^2 \underset{\cdot}{\mathrm{x}} + o(1) \\
&\Longrightarrow \quad \int_{-\infty}^{\infty} \left[\eta^F(x - \vartheta_0) - \hat{u} f(x - \vartheta_0) \right]^2 \underset{\cdot}{\mathrm{x}} \\
&= \quad \int_{-\infty}^{\infty} \left(\eta^F(y) - \hat{u} f(y) \right)^2 \underset{\cdot}{\mathrm{y}} = \delta.
\end{aligned}$$

So that the limit of the statistic Δ does not depend on ϑ_0, and the test $\Psi_T = 1\!\!1_{\{\Delta_T \geq c_\varepsilon\}}$ with c_ε the solution of

$$\mathbf{P}\left(\Delta \geq c_\varepsilon\right) = \varepsilon$$

belongs to $\mathcal{K}_\varepsilon$.

Remark. It can be shown that in the case of Kolmogorov-Smirnov tests

$$\varphi_T = 1\!\!1_{\{\omega_T > p_\varepsilon\}}, \qquad \Phi_T = 1\!\!1_{\{\Omega_T > q_\varepsilon\}}$$

where

$$\omega_T = \sup_x \left|\hat{f}_T\left(x\right) - f\left(x - \hat{\vartheta}\right)\right| \sqrt{T}, \qquad \Omega_T = \sup_x \left|\hat{F}_T\left(x\right) - F\left(x - \hat{\vartheta}\right)\right| \sqrt{T}$$

the limit distributions of these statistics (under hypothesis) do not depend on ϑ. The proofs can be done following the same lines as in Kutoyants [9] and Negri [14] respectively.

5 Consistency

In this section we discuss the consistency of the proposed tests. We study the tests statistics under the alternative hypothesis that is defined as

$$\mathcal{H}_1 : \; S(\cdot) \notin \overline{\mathcal{S}(\Theta)},$$

where $\overline{\mathcal{S}(\Theta)} = \{S\left(x - \vartheta\right), \vartheta \in [\alpha, \beta]\}$.

Under this hypothesis we have:

Proposition 5.1. *Let all drift coefficients under alternative satisfy the conditions $\mathcal{ES}$, $\mathcal{A}_0$, and $\mathcal{A}$, then for any $S(\cdot) \notin \overline{\mathcal{S}(\Theta)}$ we have*

$$\mathbf{P}_S\left(\delta_T > d_\varepsilon\right) \longrightarrow 1,$$

and

$$\mathbf{P}_S\left(\Delta_T > c_\varepsilon\right) \longrightarrow 1.$$

Proof. Remember that under hypothesis $\mathcal{H}_1$, the MLE $\hat{\vartheta}_T$ converges to the point which minimize the distance

$$D(\vartheta) = \mathbf{E}_S\left(S_*(\xi - \vartheta) - S(\xi)\right)^2,$$

where ξ is the random variable of invariant density $f_S(x)$ (See [9], Proposition 2.36):

$$\hat{\vartheta}_T \longrightarrow \hat{\vartheta}_0 = \arg\inf_{\vartheta\in\Theta} D(\vartheta).$$

In addition, denoted with $\|\cdot\|$ the norm in L^2, we have

$$\begin{aligned}\mathbf{P}_S\left(\delta_T > d_\varepsilon\right) &= \mathbf{P}_S\left(\left\|\hat{f}_T(\cdot) - f(\cdot,\hat{\vartheta}_T)\right\|^2 > d_\varepsilon\right)\\ &\geq \mathbf{P}_S\left(\left\|f_S(x) - f(x-\hat{\vartheta}_T)\right\|^2 - \left\|\hat{f}_T(x) - f_S(x)\right\|^2 > d_\varepsilon\right).\end{aligned}$$

We can deduce

$$\begin{aligned}\left\|f_S(x) - f(x-\hat{\vartheta}_T)\right\|^2 &= T\int_{-\infty}^{\infty}\left(f_S(x) - f(x-\hat{\vartheta}_T)\right)^2 \underset{\cdot}{\mathrm{x}}\\ &= T\int_{-\infty}^{\infty}\left(f_S(x) - f(x-\hat{\vartheta}_0) + o(1)\right)^2 \underset{\cdot}{\mathrm{x}}\\ &= (C+o(1))T \longrightarrow \infty, \quad \text{as } T\longrightarrow\infty.\end{aligned}$$

Moreover

$$\begin{aligned}\mathbf{E}_S\left(\left\|\hat{f}_T(x) - f_S(x)\right\|^2\right) &= \mathbf{E}_S\left(T\int_{-\infty}^{\infty}\left(\hat{f}_T(x) - f_S(x)\right)^2 \underset{\cdot}{\mathrm{x}}\right)\\ &\leq C\int_{-\infty}^{\infty}\mathbf{E}_S(\eta_T(x)^2)\underset{\cdot}{\mathrm{x}} \leq C\int_{-\infty}^{\infty} \mathrm{e}^{-2\gamma|x|}\underset{\cdot}{\mathrm{x}} < \infty.\end{aligned}$$

And finally we have the result for δ_T:

$$\mathbf{P}_S\left(\delta_T > d_\varepsilon\right) \geq \mathbf{P}_S\left(\left\|f_S(x) - f(x-\hat{\vartheta}_T)\right\|^2 - \left\|\hat{f}_T(x) - f_S(x)\right\|^2 > d_\varepsilon\right) \longrightarrow 1.$$

A similar result can be obtained for Δ_T.

6 Numerical Example

We consider the Ornstein-Uhlenbeck process. Remind that the tests for O-U process were studied in [11] as well. Suppose that the observed process under the null hypothesis is

$$\underset{\cdot}{\mathrm{X}}_t = -(X_t - \vartheta_0)\underset{\cdot}{\mathrm{t}} + \underset{\cdot}{\mathrm{W}}_t, \quad X_0,\ 0\leq t\leq T.$$

The invariant density is $f(x-\vartheta_0)$, where $f(x) = \pi^{-1/2}\mathrm{e}^{-x^2}$.

The log-likelihood ratio is

$$L(X^T,\vartheta) = -\int_0^T (X_t-\vartheta)\underset{\cdot}{\mathrm{X}}_t - \frac{1}{2}\int_0^T (X_t-\vartheta)^2\underset{\cdot}{\mathrm{t}},$$

so that the MLE $\hat{\vartheta}_T$ can be calculated as

$$\hat{\vartheta}_T = \frac{1}{T}\int_0^T X_t\underset{\cdot}{\mathrm{t}} + \frac{X_T - X_0}{T}.$$

The Fisher information in this case equals to 1, and the LTE is

$$\hat{f}_T(x) = \frac{1}{T}(|X_T-x| - |X_0-x|) - \frac{1}{T}\int_0^T \operatorname{sgn}(X_t-x)\underset{\cdot}{\mathrm{X}}_t.$$

The conditions $\mathcal{A}_0$ and $\mathcal{A}$ are fulfilled, then the statistic is convergent:

$$\delta_T = \int_{-\infty}^{\infty}\left(\hat{f}_T(x) - f(x-\hat{\vartheta}_T)\right)^2\underset{\cdot}{\mathrm{x}} \Longrightarrow \delta = \int_{-\infty}^{\infty}\zeta_1(x)^2\underset{\cdot}{\mathrm{x}},$$

where the limit process $\zeta_1(x) = \eta(x) - \hat{u}f'(x)$ can be written as

$$\zeta_1(x) = \int_{-\infty}^{\infty}\left(2f(x)\frac{\mathbb{I}_{\{y>x\}} - F(y)}{\sqrt{f(y)}} + f'(x)\sqrt{f(y)}\right)\underset{\cdot}{\mathrm{W}}(y).$$

We have a similar result for the test based on the EDF:

$$\Delta_T = \int_{-\infty}^{\infty}\left(\hat{F}_T(x) - F(x-\hat{\vartheta}_T)\right)^2\underset{\cdot}{\mathrm{x}} \Longrightarrow \Delta = \int_{-\infty}^{\infty}\left(\zeta_2(x)\right)^2\underset{\cdot}{\mathrm{x}},$$

where the limit process can be written as

$$\zeta_2(x) = \int_{-\infty}^{\infty}\left(2\frac{F(y\wedge x) - F(y)F(x)}{\sqrt{f(y)}} + f(x)\sqrt{f(y)}\right)\underset{\cdot}{\mathrm{W}}(y).$$

We simulate 10^5 trajectories of δ (resp. Δ) and calculate the empirical $1-\varepsilon$ quantiles of δ (resp. Δ). We obtain the simulated density for δ and Δ that are showed in Graphic 1. The values of the thresholds d_ε for different ε are showed in Graphic 2.

Acknowledgments

The authors are grateful to Y. Kutoyants for his suggestions and interesting discussions.

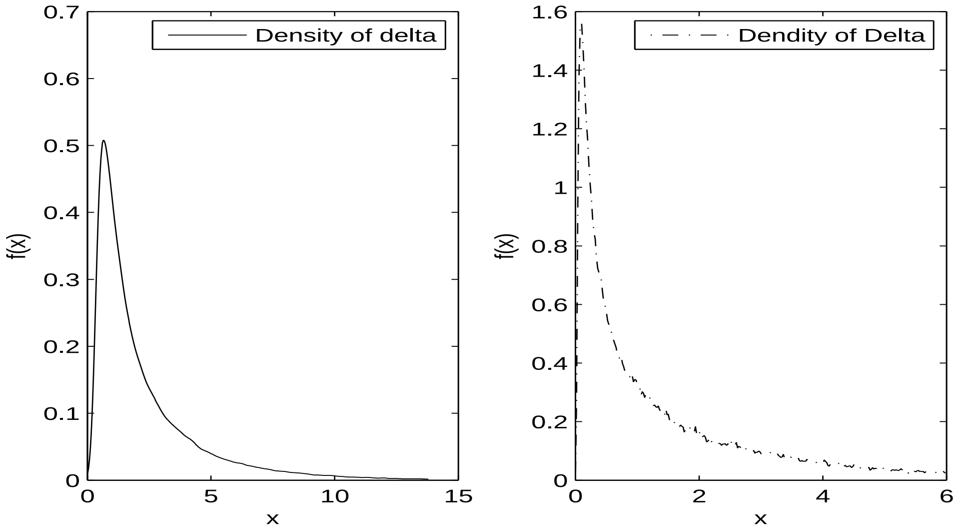

Figure 1: Density of the statistics. On the left the density of δ, on the right the density of Δ

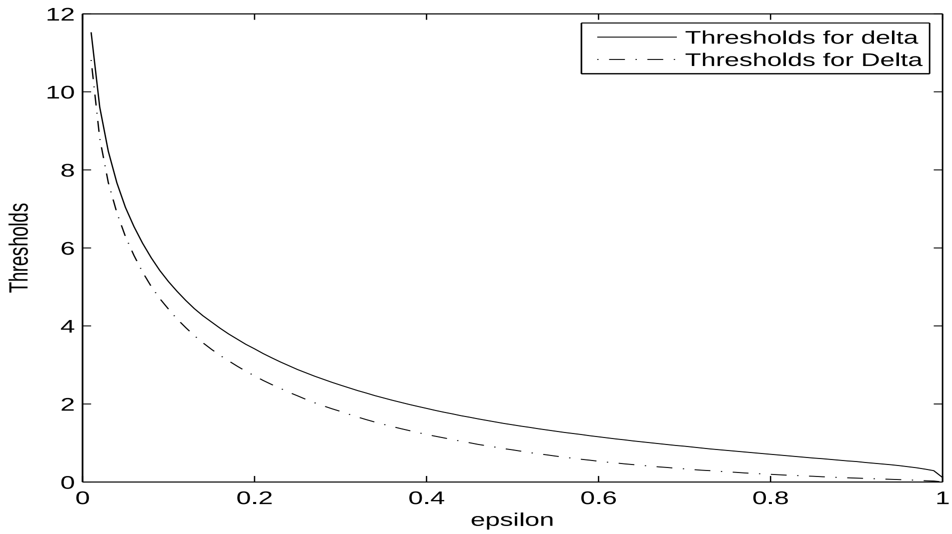

Figure 2: Threshold for different ε. The solid line represents the values for δ, the dotted line represents the values for Δ

References

[1] Babu G.J. and Rao C.R. (2004). Goodness-of-fit test when parameters are estimated. The Indian Journal of Statistics, Vol. 66, 63-74.

[2] Dachian S. and Kutoyants Y.A. (2007). On the goodness-of-fit tests for some continuous time processes. Statistical Models and Methods for Biomedical and Technical Systems, F.Vonta et al. (Eds), Birkhauser, Boston, 395-413.

[3] Darling D.A. (1995). The Cramer-Smirnov test in the parametric case. Ann. Math. Statist., 26, 1-20.

[4] Durbin J. (1973). Distribution Theory for Tests Based on the Sample Distribution Function. Society for Industrial and Applied Mathematics, Philadelphia.

[5] Durett R. (1996). Statistic Calculus: A Practical Introduction. Boca Raton: CRC Press.

[6] Fournie E. (1992). Un test de type Kolmogorov-Smirnov pour processus de diffusion ergodiques. Rapports de Recherche, 1696, INRIA, Sophia-Antipolis.

[7] Fournie E. and Kutoyants Y.A. (1993). Estimateur de la distance minimale pour des processus de diffusion ergodiques. Rapports de Recherche, 1952, INRIA, Sophia-Antipolis.

[8] Kac M., Kiefer J. and Wolfowitz J. (1995). On tests of normality and other tests of goodness-of-fit based on distance methods. Ann. Math. Statist. 26:189-211.

[9] Kutoyants Y.A. (2004). Statistical Inference for Ergodic Diffusion Processes. Springer, London.

[10] Kutoyants Y.A. (2010). On the goodness-of-fit testing for ergodic diffusion process. Journal of Nonparametric Statistic, 22, 4, 529-543.

[11] Kutoyants Y.A. (2010). Goodness-of-fit tests for Ornstein-Uhlenbeck process. arXiv:1203.3694v1, Submitted.

[12] Martynov G. V., The Omega Square Tests, Nauka, Moscow, 1979.

[13] Martynov G. V. (1992). Statistical tests based on EDF empirical process and related question. J. Soviet. Math. 61, 2195-2271.

[14] Negri, I. (1998). Stationary distribution function estimation for ergodic diffusion process, *Statistical Inference for Stochastic Processes*, **1**, 61–84.

[15] Negri, I. and Nishiyama, Y. (2009). Goodness of fit test for ergodic diffusion processes, Ann. Inst. Statist. Math., 61, 919-928.